1.Purpose of the Analysis

The purpose of this analysis is to identify auto-related crash hotspots in Wellington and evaluate predictive modeling approaches. First, hotspots will be identified based on crash rate along the road network. Next, OLS and GWR models will be applied to predict crash occurrences, using road-related features and spatial variables. The comparison between OLS and GWR will help determine which model more accurately captures geographic variations in crash risk, enabling better-targeted safety interventions.

2. Description of the Datasets

2.1 New Zealand Car Crash Report Dataset (2000–2024): This dataset provides detailed reports of car crashes across New Zealand, including Wellington. Each crash is represented as a point and contains various attributes as Table 1 below.

|  |  |
| --- | --- |
| Category | Attributes |
| Crash Details | Severity, Crash Year, Direction |
| Crash Location | Coordinates (Crash Location 1, 2), Area Unit IDs |
| Crash Outcome | Fatal Count, Serious Injury Count, Minor Injury Count |
| Road Features | Number of Lanes, Speed Limit, Road Surface Type, Intersection Involvement |
| Weather condition | Weather A (Rain, Snow), Weather B (Frost, Wind) |

This dataset forms the core of the analysis as it provides the actual crash incidents. However, since crash reports are point data, it is necessary to allocate each crash to the relevant road segment to perform meaningful analysis at the road level. The crash report's spatial component (coordinates) allows us to map the data, while the other attributes provide context for crash severity and contributing factors.

2.2. Road Section Geometry Dataset: This dataset provides the geometrical features of the road network in Wellington, such as road geometry like line segments representing each road section and road features like Road type (e.g., accessway, roadway). This dataset is critical for assigning crashes to specific road segments. In addition, it contains relevant road attributes like road type which could be useful road characteristics correlate with crash frequency. By matching crash points to the road network, this dataset allows us to compute crash rates for individual segments.

2.3. SA2 Boundary Dataset (NZ States Equivalency to TAZ): The SA2 boundaries are a geographic unit used in New Zealand for spatial aggregation, similar to TAZs in American crash studies. It provides a mid-level aggregation between smaller meshblocks and larger area units, ensuring meaningful grouping without over-segmentation and align with road characteristics. This spatial framework provides a consistent way to measure crash risk while accounting for regional differences. This provides spatial component like SA2 zone boundaries and census attribute within that boundary.

2.4 Other Data: Since the New Zealand Car Crash Dataset only have the speed limit attribute where they actual have the crash, in order to explore the overall speed limit of all the road in Wellington we will use the National Speed Limit Register (NSLR) from New Zealand Government website. Wellington’s complex terrain, with steep roads and elevation changes, requires a detailed understanding of how elevation might impact traffic crashes. To account for this, we will use the Digital Elevation Models (DEM) dataset for Wellington from New Zealand Land Information (LINZ). The DEM provides elevation values for terrain features such as mountains, valleys, and hills, which can be analysed to determine if there is a correlation between elevation and crash frequency.

3. How to access the data from Python

All datasets will be accessed and processed using Python, leveraging packages like the geopandas or rioxarray for handling spatial or raster data. The data will then be filtered to get the appropriate information.

4. Summary

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset Name | Source | Spatial Component | Purpose | Access Method |
| Crash Analysis System (CAS) Dataset | New Zealand Transport Agency (NZTA) | Geolocation (coordinates of crashes) | Detailed reports of car crashes with road characteristics and context | Publicly available via [NZTA website](https://opendata-nzta.opendata.arcgis.com/datasets/8d684f1841fa4dbea6afaefc8a1ba0fc_0/explore?filters=eyJjcmFzaFllYXIiOlsyMDEwLDIwMjRdfQ%3D%3D&location=-3.249422%2C0.000000%2C1.28&showTable=true). |
| Wellington Road Section Geometry | Land Information New Zealand (LINZ) | Geolocation (road segments and classifications) | Allocate crash to the relevant road segment | Retrieved from [LINZ](https://data.linz.govt.nz/layer/53378-nz-roads-road-section-geometry/) spatial data infrastructure. |
| Population Data by Statistical Area 2 (SA2) | Statistics New Zealand | SA2 boundaries (mapped population areas) | Further road segmentation and combine population density attribute | Publicly available from [Statistics New Zealand](https://datafinder.stats.govt.nz/layer/105008-estimated-resident-population-at-30-june-2018-by-statistical-area-2/). |
| National Speed Limit Register (NSLR) | New Zealand Government | Geolocation (road segments and speed limits) | Analyse the influence of speed limits across Wellington's road network. | Publicly available from [New Zealand government data](https://catalogue.data.govt.nz/dataset/national-speed-limit-register-nslr1) |
| Digital Elevation Models (DEM) Dataset | Land Information New Zealand (LINZ) | Elevation (raster) data of terrain (mountains, valleys, hills) | Analyse the effect of terrain on traffic crash frequency. | Retrieved from  [LINZ](https://data.linz.govt.nz/layer/53621-wellington-lidar-1m-dem-2013-2014/) spatial data infrastructure. |